GLOBAL JOURNAL OF MATHEMATICAL SCIENNCES VOL. 6, N0. 1, 2007: 53 - 56
COPYRIGHT (C) BACHUDO SCIENCE CO. LTD. PRINTED IN NIGERIA. ISSN 1596-6208

APPLICATION OF SINGH ET AL UNBIASED ESTIMATOR IN A DUAL TO RATIO-CUM-PRODUCT ESTIMATOR IN SAMPLE SURVEYS TO DOUBLE SAMPLING DESIGN

A.A. ADEWARA

(Received 8 November 2006. Revision Accepted 17 May 2007)

ABSTRACT

This paper applied an unbiased estimator in a dual to ratio–cum-product estimator in sample surveys to double sampling design. Its efficiency over the conventional biased double sampling design estimator was determined based on the conditions attached to its supremacy. Three different data sets were used to testify to when our conventional biased double sampling design estimator would be preferred to this alternative unbiased type. A data set was also used to show when this conventional biased double sampling design estimator would not be preferred but instead, the alternative unbiased double sampling design estimator

KEYWORDS: Unbiased Estimator, Ratio-cum-product estimator, Double sampling design, efficiency, differences

1. INTRODUCTION

If the information needed to improve on the estimate of the character under study is lacking, and if it is convenient and cheap to do so, then information on the auxiliary variable is collected from a preliminary large sample (n') (Okafor, 2002). While information on the variable of interest, y, is collected from a second sample (n), may be, a sample of the preliminary sample or may be an independent sample selected from the entire population. When the second sample (n) is independent of the preliminary sample,(n'), ($n' > n$), information on both the auxiliary and the main character is obtained from the second sample(Okafor, 2002).

The double sampling design ratio estimator of the population mean of y is given by, $\bar{y}_{ds1} = \frac{\bar{y}}{\bar{x}}\bar{x}'$, where $\bar{x}' = \frac{1}{n'}\sum_{i=1}^{n'} x_i$ is the sample estimate of $\bar{X}$ obtained from the first phase sample, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ and $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ are the sample estimates of $\bar{Y}$ and $\bar{X}$ respectively, obtained from the second phase sample (Cochran,1977; Okafor,2002 and Singh et al,2005).

Being a biased double sampling design estimator, (Cochran (1977), Raj (1972), and Raj and Chandhok (1999)),

$$bias(\bar{y}_{ds1}) = \frac{1}{\bar{X}}(\frac{n'-n}{nn'})(RS^2_x - S_{xy}) \tag{1}$$

and

$$mse(\bar{y}_{ds1}) = (\frac{N-n'}{Nn'})S^2_y + (\frac{n'-n}{nn'})(S^2_y - 2RS_{xy} + R^2S^2_x) \tag{2}$$

where,

$S^2_x = \sum_{i=1}^{N}(x_i - \bar{X})^2/(N-1)$, $S^2_y = \sum_{i=1}^{N}(y_i - \bar{Y})^2/(N-1)$,

$S^2_{xy} = \sum_{i=1}^{N}\{(x_i - \bar{X})(y_i - \bar{Y})\}/(N-1)$ and $R = \frac{\bar{Y}}{\bar{X}}$

In this paper, we applied Singh et al (2005) unbiased estimator which utilizes the work of Bangyopadhyay (1980) and Srivenkataramana (1980) in a dual to ratio-cum-product estimator in sample surveys to double sampling design.

Let $x^*_i = (1+g)\bar{X} - g\bar{x}, i = 1,2,\ldots,N$, where, $g = \frac{n}{N-n}$. Then clearly, $\bar{x}^* = (1+g)\bar{X} - g\bar{x}$ is an unbiased estimator of $\bar{X}$ and $corr(\bar{y},\bar{x}^*) = -\rho_{xy}$ {Singh et al (2005)}.

We now apply this suggested unbiased estimator $\bar{x}^*$ to double sampling design (instead of a dual to ratio-cum-

A.A Adewara, Department of Statistics, University of Ilorin, Ilorin, Kwara State. Nigeria

product estimator that Singh et al (2005) applied this for) as $\bar{y}_{ds2} = \frac{\bar{y}}{\bar{x}^*}\bar{x}'$ where, $\bar{x}^* = (1+g)\bar{X} - g\bar{x}$, $\bar{y}$ and $\bar{x}'$ are as in section 1.

Being an unbiased double sampling design estimator, Singh et al (2005),

$$bias(\bar{y}_{ds2}) = 0 \qquad (3)$$

and

$$mse(\bar{y}_{ds2}) = (\frac{N-n'}{Nn'})S^2_y + (\frac{n'-n}{nn'})(S^2_y - 2gRS_{xy} + R^2g^2S^2_x) \qquad (4)$$

2. COMPARISON

(a) Efficiency of $\bar{y}_{ds2}$ over $\bar{y}_{ds1}$

The concept of efficiency in statistical estimation is due to Fisher (1921) and is an attempt to measure objectively the relative merits of several possible estimators which was that of the estimates. An estimator is being regarded as more "efficient" than another if it has smaller estimate (Kendall and Buckland, 1982).

Therefore from (2) and (4) above, our unbiased double sampling design estimator, $\bar{y}_{ds2}$ will be better and more efficient than the conventional biased double sampling design estimator, $\bar{y}_{ds1}$ whenever $mse(\bar{y}_{ds2})$ < $mse(\bar{y}_{ds1})$.

That is,

$$(S^2_y - 2gRS_{xy} + R^2g^2S^2_x) < (S^2_y - 2RS_{xy} + R^2S^2_x)$$

But one can see from (4) that whenever $g = \frac{n}{N-n} = 1$, eq.(2) = eq.(4) , meaning that both estimators are equal.

Here g is a function of both the sample (n) and the population sizes (N), so g is not arbitrarily chosen.

b. Differences

(i) While $\bar{y}_{ds1}$ is a biased double sampling design estimator, $\bar{y}_{ds2}$ is an unbiased double sampling design estimator

(ii) In eq(2), the coefficients of S_{xy} and S^2_x are -2R and $+R^2$ respectively but in eq. (4), the coefficients of S_{xy} and S^2_x are -2gR and $+g^2R^2$ respectively. So, their differences lies on the coefficients of S_{xy} and S^2_x.

3. DATA FOR VERIFICATION OF HYPOTHESES.

Four data sets are used in this paper to establish the efficiency of these estimators called populations 1 – 4 as shown below:-.

Population 1:- Okafor (2002), p 269 - 272. From a total of 912 villages, a simple random sample and subsample of 100 villages and 30 villages are chosen respectively to obtain the total area in hectares (x) and the area with cassava (y). Here,

x: village Area, y: Area under cassava.

N = 912, n'=100, n = 30, $\bar{x}'$=1402 42, $\bar{x}$ = 1372.614, $\bar{y}$=733.347, R = 0 53427, S^2_x = 791376.06, S^2_y =151031.29, S_{xy} = 268108.86

Population 2:- Okafor (2002), p. 389 - 390. From a total of 481 cooperative farmers in 3 rural areas, a simple random sample and subsample of 134 and 28 registered cooperative farmers are chosen respectively Here,

x: Household size, y: Local Area planted,

N = 481, n'=134, n = 28, $\bar{x}'$=5, $\bar{x}$ = 7.2857, $\bar{y}$ = 6.6946, R = 0.9189, S^2_x = 7.6190, S^2_y = 18.2101, S_{xy} = 4.6261

Population 3:- Raj (1972), p. 89 – 91. In order to estimate the total cultivation area in a commune containing N = 850 parcels of land, a random sample and subsample of 100 and 30 parcels are selected and eye estimates x of cultivated area and each parcel measured for cultivated area y are obtained. Here, x: Eye estimate of cultivated area, y: Each parcel measured for cultivated area.

N = 850, n'=100, n = 30, $\bar{x}'$=4 31, $\bar{x}$ = 3.93, $\bar{y}$ = 4.25, R=1.0814,

S^2_x = 9.03, S^2_y = 9 0274, S_{xy} = 8.6353

Population 4:- Raj (1972), p. 102 - 103. From a directory listing (N) 3,500 large manufacturing establishment, a random sample and subsample of 158 and 30 establishments were taken and questionnaires mailed to obtain information on the number of paid employees x and more accurate data on employment y Here.

x: Reported figures on employment, y: True figures on employment,
$N = 3500$, $n' = 158$, $n = 30$, $\bar{x}' = 46.99$, $\bar{x}$ 47.87, $\bar{y}$ 50.97, $R = 1.0648$,
S^2_x = 850.0506, S^2_y 870.6551, S_{xy} 97.846

RESULTS

The results obtained on, $(S^2_y - 2RS_{xy} + R^2S^2_x)$, $(S^2_y - 2gRS_{xy} + R^2g^2S^2_x)$, g, $|bias(\bar{y}_{ds1})|$, $|bias(\bar{y}_{ds2})|$, $mse(\bar{y}_{ds1})$ and $mse(\bar{y}_{ds2})$ using the above information for the four populations considered are presented below in Table 1:-

Table 1:- mse and bias of $\bar{y}_{ds1}$ and $\bar{y}_{ds2}$

	Pop 1	pop 2	pop 3	pop 4
$(S^2_y - 2RS_{xy} + R^2S^2_x)$	90,440.13	16.1416	0.9109	1626.069
$(S^2_y - 2gRS_{xy} + R^2g^2S^2_x)$	141,548.30	17.7092	8.3583	868.9257
g	0 0340	0.0618	0.0366	0.0086
$\lvert bias(\bar{y}_{ds1})\rvert$	4.5576	0.0179	0.0513	0.0552
$\lvert bias(\bar{y}_{ds2})\rvert$	0	0	0	0
$mse(\bar{y}_{ds1})$	3454 978	0.5541	0.1009	49.1724
$mse(\bar{y}_{ds2})$	4647.501	0.5984	0.2747	28.7264

5. DISCUSSIONS

From the results shown in Table 1 above, we observed that, for populations 1 – 3, $(S^2_y - 2gRS_{xy} + R^2g^2S^2_x) > (S^2_y - 2RS_{xy} + R^2S^2_x)$ indicating that the conventional double sampling design estimator, $\bar{y}_{ds1}$ still maintains its supremacy over the unbiased double sampling design estimator, $\bar{y}_{ds2}$. But for population 4, $(S^2_y - 2gRS_{xy} + R^2g^2S^2_x) < (S^2_y - 2RS_{xy} + R^2S^2_x)$, meaning that our unbiased double sampling design estimator, $\bar{y}_{ds2}$ is better than the conventional double sampling design estimator, $\bar{y}_{ds1}$.

6. CONCLUSION

We conclude that, our unbiased double sampling design estimator, $\bar{y}_{ds2}$, will be preferred to the biased conventional double sampling design estimator, $\bar{y}_{ds1}$, whenever,

(i) $g = \dfrac{n}{N-n} < 1$ and

(ii) $(S^2_y - 2gRS_{xy} + R^2g^2S^2_x) < (S^2_y - 2RS_{xy} + R^2S^2_x)$

ACKNOWLEDGEMENT

The author is grateful to Prof. G.A. Olayi of Department of Mathematics, University of Calabar, Calabar, Nigeria for helpful comments and suggestions which have enabled me to correct a number of errors and also to improve the presentation of the manuscript.

REFERENCES

Bandyopadhyay, S., 1980. Improved ratio and product estimators, Sankhya Series C 42, 45 - 49

Cochran, W.G., 1977. Sampling Techniques, 3rd Edition. John Wiley, New York.

Fisher, R. A., 1921. On the mathematical foundations of the oretical statistics. Phil. Trans. Roy Soc., London, A222.

Kendall, M.G. and Buckland, W.R., 1982. A dictionary of Statistical terms, 4th edition revised and enlarged. Longman Group Ltd, USA.

Okafor, F.C., 2002. Sample Survey Theory with Application. Afro-Orbis Publications Ltd, University of Nigeria. Nsukka.

Raj, D., 1972. The Design of Sample Surveys. Mc Graw-Hill Inc. USA

Raj, D and Chandhok, P., 1999. Sample Survey Theory. Narosa Pub. House, London

Singh, H. P., Singh, R., Espejo, M. R. and Pineda, M. D., 2005. On the efficiency of a dual to ratio – cum – product estimator in sample surveys. Mathematical Proceedings of the Royal Irish Academy, 105 A (2), 51 - 56

Srivenkataramana, T., 1980. A dual to ratio estimator in sample surveys. Biometrika 67, 199 - 204